

GRIPS Discussion Paper 16-23

**国会審議映像検索システムとそのニュース動画への応用・検証**  
**Video Retrieval System for Diet Deliberations and its Applicability to**  
**News Videos**

**増山 幹高**

**Mikitaka Masuyama**

November 2016



**GRIPS**

NATIONAL GRADUATE INSTITUTE  
FOR POLICY STUDIES

National Graduate Institute for Policy Studies  
7-22-1 Roppongi, Minato-ku,  
Tokyo, Japan 106-8677

## 国会審議映像検索システムとそのニュース動画への応用・検証

増山幹高

政策研究大学院大学

### 要旨

国会審議映像検索システムは、国会会議録の文字情報から、発言に対応する映像をピンポイントで再生するものであり、審議動画を音声認識することによって、会議録と審議映像を同期させている。これにより、国会会議録がキーワード検索できるように、国会審議映像も発言内容から検索できるようになり、審議映像の部分再生、字幕付与が可能となる。審議動画は衆参両院事務局のサーバー上で再生し、審議映像の URL をインターネットで共有することも容易である。

国会審議には会議録に含まれない様々な情報がある。国会審議映像検索システムを活用すれば、パネルなどの視覚的資料も瞬時に確認することができるし、会議録の公開を待たずとも音声認識によって映像情報と文字情報を同刻する技術を審議映像の検索に活用することもできる。また、音声認識は審議映像に限らず、ニュース動画にも適用することが可能であり、本稿では、国会審議映像の部分的視聴を可能にする「国会審議映像検索システム」を概説したうえで、政治関連のニュース動画に音声認識を適用する可能性を検証する。

### 付記

本稿は「政策情報のユニバーサル化・国際化に関する実証と実践」(2015 年度～2019 年度・課題番号 15H05727)、GRIPS 政策研究センター「国会アーカイブス・プロジェクト」(2015 年度)による部分的成果である。

## Video Retrieval System for Diet Deliberations and its Applicability to News Videos

### Summary

This paper offers an overview of the video retrieval system we have developed for the Japanese Diet, and examines how we can make political news videos searchable with keywords by using the sound recognition techniques to match up the Diet proceedings and deliberation videos. With our video retrieval system one can directly retrieve the moment of video feeds he or she is interested in, visually understand the flow of parliamentary debate, and check the facial expression and body language of the speaker. In this paper, we briefly describe how our video retrieval system works, report a preliminary analysis on the speech recognition results of news videos, and suggest alternative ways to utilize visual information to better understand political discussion both inside and outside the Diet.

## 1. はじめに

比較議会情報プロジェクトでは<sup>1</sup>, 文字情報に偏ってきた国会審議や政治報道に対する従来の分析から脱し, 音声や画像, 映像を活用することによって, これまで捨象されてきた立法情報の異次元を解明することに取り組んでいる.

国会の会議録が重要な立法情報であることに疑いはないが, 国会で起きていることが全て会議録に残るわけではない. 国会の記録とはなにか. 憲法第 57 条は国会に会議の記録を保存し, 公表・頒布することを求めている. 記録技術の限られた時代において, 紙媒体に記述された文字情報は唯一の記録であり, 会議録は帝国議会の最初から 1 世紀余に連綿と残されている. 国会は比較的発言を忠実に記録しているが, 会議録は発言を 100%再現しているわけではない. 発言中に「,」や「.」と口にするわけでもなく, 「えー」「そのー」は会議録には残っていない. ケバという無駄な言葉を取り除き, 話し言葉を書き言葉に直す整文が行われている.

つまり, 会議録は公的な記録として貴重な情報源であるが, 何らかの加工を施された文字情報であり, 様々な非文字情報が捨象されていることに留意する必要がある. 帝国議会が始まった頃の様子は, 日本史の教科書に載っているような錦絵と言われる版面から推測するしかないが, 写真や録音の技術が進歩・普及し, 国会のラジオやテレビによる中継も 1950 年代には始まり, 国会議事堂を訪れ, 会議を傍聴しなくとも, 国会審議はお茶の間に届けられ, 国会の様子はより視覚的・聴覚的に把握できるようになっている.

議会の歴史とは, 世の東西を問わず, 議会と国民の間にある空間的, 時間的な隔たりを取り払っていく過程であるとも言える. 演説会における雄弁さを政治家がアピールした時代と異なり, マスメディアが発達することによって政治家は否応なしに日常的に業績を誇示することを迫られるようになる. 議員がテレビ中継される予算委員会での質疑に立ちたがり, 答弁を求める閣僚ではなく, カメラのほうに大きなボードを向けたりするのは, テレビを通じた国民の視線を意識しているからに他ならない.

インターネットによる動画配信は議会と国民の時空間をさらに短縮する. 国会では, 1990 年代の終わりから衆参両院事務局が審議映像をインターネットで配信し, ライブストリーミングでの視聴やライブラリに保存された審議動画がいつでもどこでも視聴できる.

<http://www.shugiintv.go.jp/index.php>

<http://www.webtv.sangiin.go.jp/webtv/index.php>

比較議会情報プロジェクトは, こうした時空間的制約のさらなる解消を目指す試みとして, 国会図書館の提供する会議録と衆参両院の事務局が配信する審議動画をリンクさせ, 発言のキーワード検索から審議映像をピンポイントで再生することを目指している. 具体的には, 音声認識によって映像情報と文字情報を同刻することにより, 審議映像を発言単位でキーワード検索し, 審議動画を配信元のサーバー上で再生しつつ, 字幕をタイムラグなしに付すことを可能にしている. こうした試みにより, 国会の審議映像の利用方法が革新的に改善され, 国会審議に関する文字情

---

<sup>1</sup> <http://www3.grips.ac.jp/~clip/>

報と映像情報を体系的に検証することが期待され、審議動画が視覚障害者や聴覚障害者にも活用されるようになるとともに、音声認識による動画検索を地方議会や審議会などの会議全般、ニュースの動画配信に応用する可能性も検証できるようになる。

本稿では、「国会審議映像検索システム」を概説し、その応用可能性を検証する試みとして、ニュース動画の音声認識に関する分析を報告する<sup>2</sup>。

## 2. 国会審議映像検索システム

国会審議について議員の発言内容に対応する審議映像を検索し、該当する審議映像の部分的視聴を可能にする「国会審議映像検索システム」は、2012年11月26日から一般公開され、2014年4月18日からは自動翻訳機能を活用した英文入力サイトが運用されている<sup>3</sup>。

<http://gclip1.grips.ac.jp/video/>

国会審議映像検索システムは、国会図書館の提供する国会会議録検索システムで発言内容を検索できるように、会議録の文字情報から、発言に対応する映像にピンポイントで到達することを目指している。これにより、例えば、ニュースや新聞記事で「〇〇議員が△△と国会審議で発言」という情報に接した際、「〇〇 △△」でキーワード検索し、その瞬間の審議映像をピンポイントで再生できるようにし、会議録からは分からない発言者の表情や臨場感、会議の流れなどが把握できるようになる。また、審議映像に字幕を付すことにより、聴覚障害者や視覚障害者が審議映像を利用することも可能になり、発言の瞬間をURLとして表現することにより、SNSを通じて審議映像をインターネットで共有することも容易になる<sup>4</sup>。

具体的には、審議動画の発言内容を検索するため、検索インデックスとして、動画中の発言のテキスト・データと、そのテキスト・データと動画中の映像シーンの同期情報が必要となる（図1）。国会審議の場合、従来は速記により、衆議院では現在は音声認識により、国会会議録が作成されており、国会審議映像検索システムでは、国会会議録のテキスト・データと審議動画の音声データの同期データを作成することにより、審議動画中の発言内容を検索することが可能になっている。

国会審議映像と国会会議録テキストの同期データの作成には、京都大学で開発した「字幕自動生成システム」を利用している<sup>5</sup>。字幕自動生成システムでは、音声ストリーム・データとテキスト・データを付与すると、それらの同期情報が生成され、簡単な操作によって数時間にわたる審議動画とテキスト・データの同期情報の作成が可能となり、審議映像の検索インデックスを作成

---

<sup>2</sup> 本稿は増山（2016）を修正加筆したものである。

<sup>3</sup> 国会審議映像検索システムについては、これまでも国内外の学会や専門誌で紹介している。Masuyama (2012), 増山・竹田 (2012), Masuyama and Takeda (2014), 鈴木他 (2014), 増山・竹田 (2015), Masuyama (2016)。

<sup>4</sup> 参議院の審議映像は、発言内容で検索可能であり、映像の部分再生も可能であるが、映像に対応する会議録を確認することはできず、字幕も付されていない。

<sup>5</sup> 河原 (2012), 秋田・河原 (2013)。

することができる。

図 1 審議動画と国会会議録の同期情報

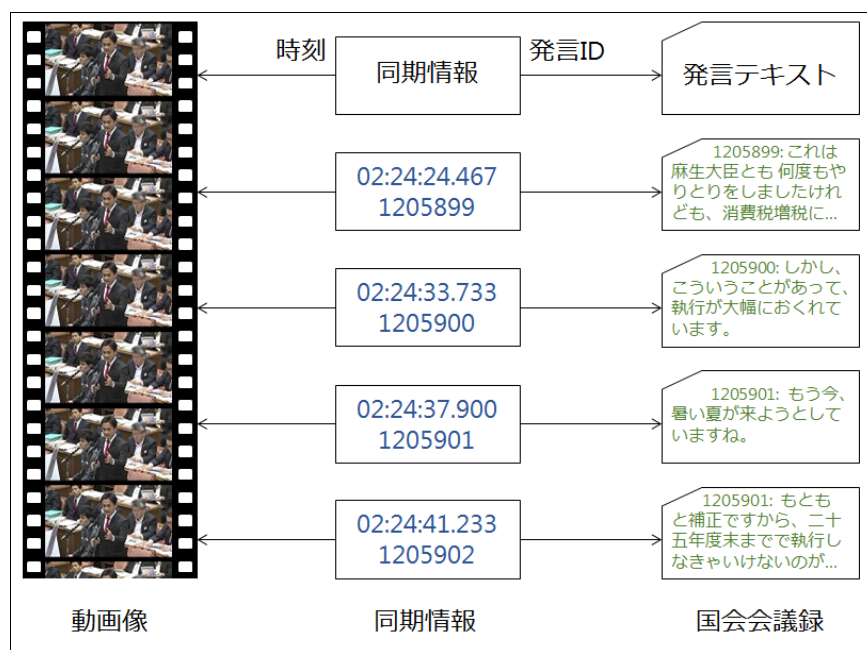


図 2 国会審議映像検索システムの構成

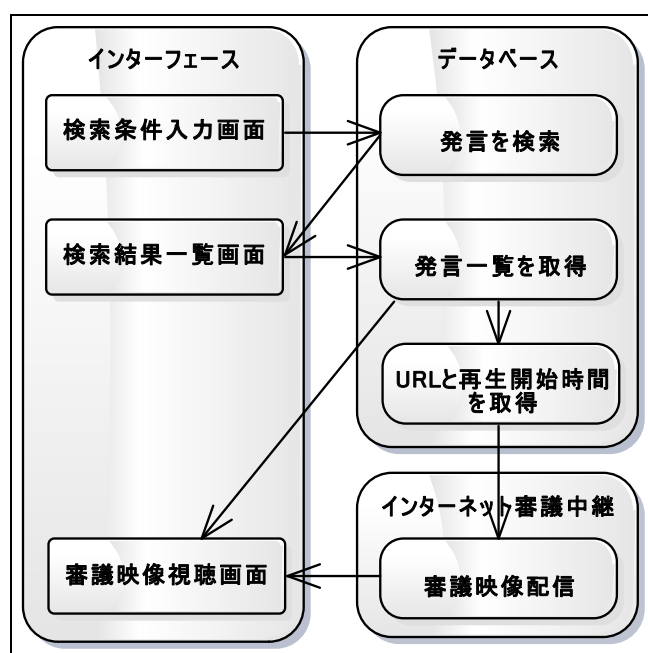


図 2 に示すように、国会審議映像検索システムは、インターフェースとデータベースの二つから構成される。データベースには、国会会議録のテキスト・データと国会会議録中の発言と審議映像の同期データが含まれる。国会会議録のテキスト・データから全文検索用のインデックスが生成され、それらが利用者の入力したキーワードに対応する発言を検索する際に利用される。検索インデックスには、会議名や発言者名もメタデータとして付与されており、これらの組み合わ

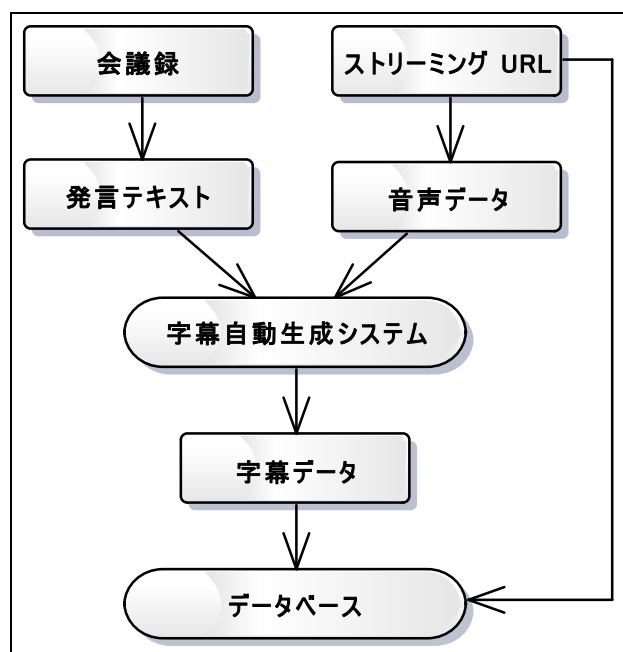
せによる検索を可能にする。国会会議録のテキスト・データは、審議映像の再生中の字幕表示や会議録の表示にも利用される。

国会審議における発言と映像の同期データは、検索結果に対応するシーンの視聴ページに遷移する際に利用される。このデータを生成・保持することにより、発言に対応する映像の冒頭からピンポイントで視聴することが可能となる。また、この同期データは、審議映像の再生中に表示する字幕の切替えや、会議録の該当発言を表示する切り替えにも利用されている。なお国会審議の動画自体は、データベースでは保持せず、配信元である衆参両院事務局のウェブサイトで再生している。つまり、国会審議映像検索システムは、そうした審議映像データのプロキシとして機能している。

また、図3は国会審議映像と国会会議録の同期データ作成の流れを示している。同期データを作成するには、国会会議録のテキスト・データと国会審議映像の音声データが必要となる。まず国会図書館の国会会議録検索システムから、国会会議録のテキスト・データを取得する。国会審議映像検索システムでは、会議録取得の一連の作業をシステム上で行う機能を開発し、取得したデータは字幕生成の処理ができる書式に変更される。

審議映像については、衆参両院事務局のウェブサイトから会議録に対応する審議映像のストリーミング・データを取得し、音声データを抽出して音声ファイルとして保存する。国会審議映像検索システムでは、審議動画のURLを特定すると、動画のストリーミング・データを取得し、音声ファイルとして保存する処理が自動的に行われる。このように国会会議録と音声ファイルを抽出し、京都大学の字幕自動生成システムを用いて同期データの作成を行い、国会会議録のテキスト・データおよびストリーミング・データのURLと合わせて、発言と映像の同期データが国会審議映像検索システムに登録される。

図3 発言と映像の同期データ作成の流れ

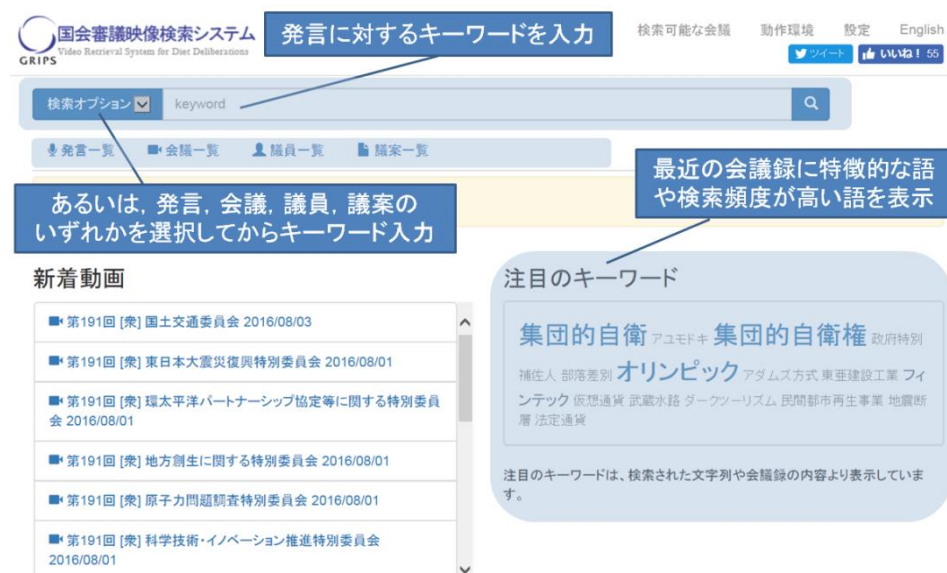


2016年4月にリニューアルした国会審議映像検索システムのトップページでは、上部の検索



キーワード入力欄に検索条件を入力するか、検索オプションから発言、会議、議員、議案のいずれかを選択し、検索条件を入力することで国会審議映像を検索することができる(図4)。例えば、検索オプションで「議員」を選んだ場合、「安倍 憲法」と入力すると、発言者としての「安倍」の発言中の「憲法」を優先的に検索するが、検索オプションを選択しない場合、誰かの発言中の「安倍」と「憲法」も同様に検索される。

図4 国会審議映像検索システムのトップページ



「注目のキーワード」欄には、新たに集中的に発言され、検索される単語の上位15位が表示されている。こうしたキーワードは、検索可能な審議映像に対応する国会会議録から、会議録に頻出する単語であり、最新のものほどウェイトをかけ、また特定の会議のなかで頻度の高いものによりウェイトをかけるように抽出されるとともに、検索頻度の高さも加味している<sup>6</sup>。また、利用者側の注目度を反映する仕組みとして、実際に検索入力されるものほど大きく表示している。これらのキーワードのいずれかをクリックすることで、そのキーワードを入力して検索することとなり、続けて検索条件を追加指定することも可能である。

図5は検索条件に「憲法改正」と入力し、2015年9～12月の衆議院の会議に絞り込んだ検索結果を示している。入力したキーワードを含む発言の一覧が右側の検索結果エリアに最近のものから表示され、個々の検索結果には、タイトル部分に、会議名、開会日が表示され、発言者名に続いて、検索キーワードが含まれる発言部分の会議録が表示される。タイトル部分は審議映像視聴ページへのリンクとなっており、ここをクリックすることで映像を視聴することができる。また、左側のフィルタエリアには、最初に入力した検索条件に加えて、開会日、衆参のいずれか、

<sup>6</sup> 具体的には、 $s(word) = w \frac{f(word)}{\max(f(word))} + (1-w) \frac{q(word)}{\max(q(word))}$  による。  $f(word)$  は単語の

会議録における頻度であり、時間的に最新であるほど、また特定の会議中における頻度の高さでウェイトをかけている。  $q(word)$  は単語の検索頻度である。  $w$  はこれら両頻度のウェイトであり、現時点では0.5を想定している。

発言の出現回数の多い発言者 5 名および 10 会議による絞り込みメニューが表示される。

図 5 「憲法改正」を入力して検索した結果（衆議院 2015 年 9~12 月）

The screenshot shows the GRIPS search interface. The search term '憲法改正' (Constitutional Amendment) is entered. The results are filtered for the House of Representatives (衆議院) in November 2015. A list of search results is displayed, including the date and session number. A blue box highlights the '検索結果エリア' (Search Results Area) and the '絞り込みメニュー' (Filter Menu). The filter menu includes options for '開会日' (Opening Date), '衆参' (House of Representatives/Senate), '発言者' (Speaker), '会議' (Session), and '回数' (Number of Occurrences). The search results list shows the date and session number for each result.

検索結果のいずれかをクリックすると、国会審議映像の視聴ページに遷移する（図 6）。この視聴ページに遷移すると、選択した発言の部分再生が始まる<sup>7</sup>。左側上部の動画再生エリアには審議映像が再生され、その下の字幕エリアに発言者の氏名と発言内容の字幕、再生操作のコントロールが表示される。右側の発言リストには、再生している会議の会議録全体が発言者毎に表示され、部分再生しているシーンに該当する発言の箇所が反転表示される。検索した審議映像の部分再生が終了すると、自動的に続く発言の審議映像を部分再生し、映像の再生の進行に合わせて、会議録の反転表示している箇所も変化する（1 分間ないし 3 発言まで自動的に再生する初期設定になっている）。また、右側の発言リストのいずれかをダブルクリックすると、自動的に部分再生する審議映像を切り替えることができ、検索結果の発言は黄色くハイライトされて残る。動画再生エリアの下部に部分再生している発言のシーンに対応する審議動画の URL を表示しており、その URL をツイートできるようにしてある。部分再生している動画をツイートしたい場合、再生中にツイートするためのボタンをクリックし、発言と URL をツイートすることができ、注目する発言のシーンをインターネットで共有することができる。さらに、視聴ページ下部には、発言者のプロフィールを表示し、その下には会議で審議されている議案と参加している議員名のリストが表示される（図 6 には表示されていない）。

<sup>7</sup> ただし、衆参両院事務局のウェブサイトで配信されている動画を部分再生するソフトウェアとして Microsoft Silver light が必要である。インストールされていない場合、動画配信リンクをクリックすると、自動的に Silver light のインストール承認が始まる。



図6 審議映像の視聴ページ



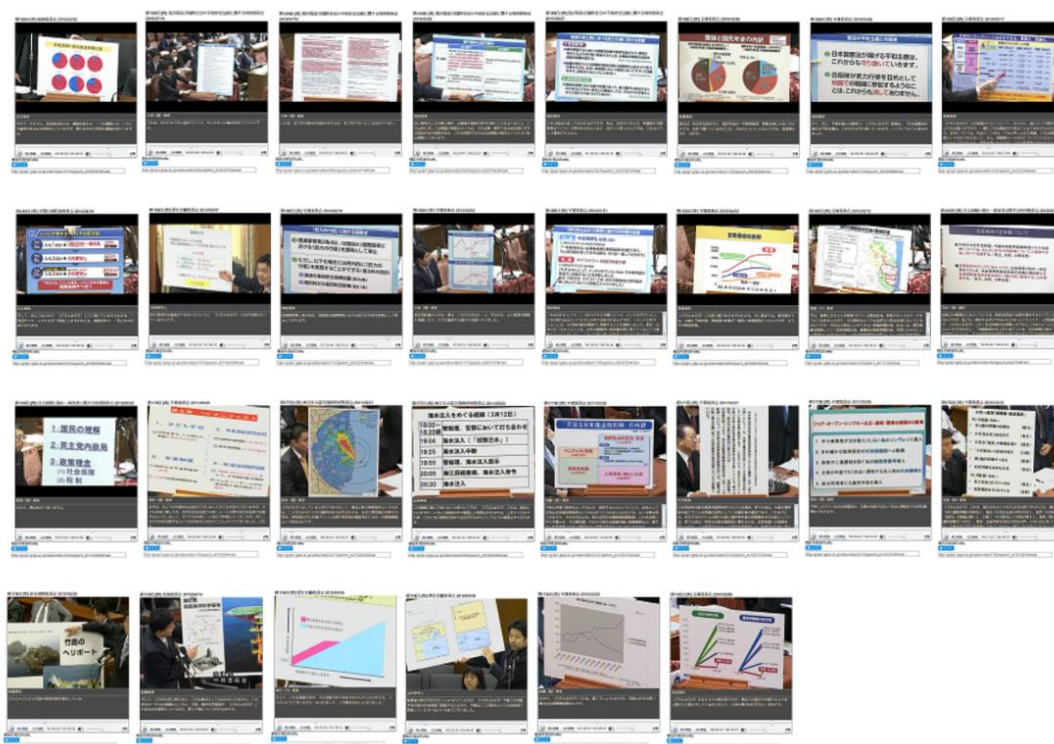
国会審議映像検索システムは、国会議員の発言内容をピンポイントで再現しようとするものであり、様々な活用法を考えることができる。図6のように、安倍首相の憲法改正についての発言を部分的に再生することによって、会議録の文字情報としてだけではなく、実際の会議における肉声として、顔の表情や声の調子、会場の雰囲気や会議の流れを併せて理解することができる。

発言の瞬間をURLとして表現しているため、SNSを通じて審議映像をインターネットで共有することも容易であり、例えば、インターネットで配信される国会関連のニュースについて、国会審議映像検索システムのURL表示機能を活用し、関連審議動画のURLを記すだけで、写真を加えたり、独自に国会審議映像を加工し、再生可能な状態で配信する以上の視覚的効果が期待される。

国会審議には会議録に含まれない様々な情報がある。例えば、委員会審議で頻繁に用いられる参考資料は、質疑に立っている議員が特に求めない限り、会議録には含まれられず、国会事務局や国会図書館で保管されているだけで、存在さえ国会関係者でなければ知る由もない。しかし、そうした参考資料は、国会審議の要点を示す視覚的情報であることが多く、審議映像に残っていれば、確認することもできる。例えば、衆議院の会議録には、(パネルを示す)というト書きがある。国会審議映像検索システムで検索したところ、34の発言において(パネルを示す)を含むことがわかった。各審議映像を部分再生すると、概ね10秒ほどでパネルにカメラがフォーカスしており、図7に整理するように、委員会審議で用いられたパネルを判別することができる<sup>8</sup>。

<sup>8</sup> <http://www3.grips.ac.jp/~clip/panel/>

図7 委員会審議におけるパネルの抽出



国会審議映像検索システムでは、審議映像に付す字幕は国会図書館の国会会議録検索システムから取得した会議録のテキスト・データに拠っており、会議録の公開までに現状で2～3週間程度かかることから、会議録に基づいて審議映像が検索できるようになり、字幕を付した形での審議映像の再生が可能となるまでに1か月程の期間が必要となる。ツイートやニュース記事などでの引用といった速報的な活用法を想定すると、よりタイムリーに審議映像を検索できることが望ましい。審議映像は翌日にはライブラリにおいて視聴可能となるため、国会審議映像検索システムでは、2015年9月以降、主要な会議については、会議録が公開されるまで、音声認識版の「会議録」で審議映像を検索可能にしている<sup>9</sup>。

### 3. ニュース動画の音声認識

音声認識によって映像情報と文字情報を同刻する技術は、地方議会や審議会などの会議全般、ニュース報道にも応用可能である<sup>10</sup>。審議映像の音声認識版による文字情報との同刻は、会議録といった「正解」のない映像検索の第一歩であり、審議映像以外の映像検索に道を開くものである。この節では、インターネットによる政治関連のニュース動画の音声認識と、それに基づく映像検索の可能性について報告する。

<sup>9</sup> 音声認識版による審議映像検索の場合、視聴ページの動画再生エリア上部に示す会議名に音声認識版によることを明記し、字幕付与機能は停止させている。

<sup>10</sup> 「地方議会審議映像検索システム」（非公開）も試行的に運用しており、現在、北見（北海道）、宮古（岩手）、筑西（茨城）、練馬（東京）、燕（新潟）、羽島（岐阜）、亀岡（京都）、瀬戸内（岡山）、鳴門（徳島）、北九州（福岡）の審議映像が検索できる。<http://gclip1.grips.ac.jp/local-assembly/>

具体的には、ニュース動画として、YAHOO!JAPAN ニュースの「映像」から主に「政治」のニュースを選択した。

<http://headlines.yahoo.co.jp/videonews/>

ニュース動画の配信元は日本テレビ系（NNN）、テレビ朝日系（ANN）、TBS 系（JNN）、フジテレビ系（FNN）であり、「映像」ページにアクセスした時点で最新の政治ニュースを選んだ。例えば、2016 年 5 月 12 日 17 時 22 分には、最新の政治関連のニュース動画は以下の通りであり、4 動画を音声認識の対象とする場合、配信時間の最新のものから 4 動画を採用した。

NNN	2016/5/12 16:05	ヘイトスピーチ解消法案 参院法務委で可決
	2016/5/11 21:45	USJ 沖縄の建設計画撤回 菅長官「残念」
	2016/5/11 17:55	米大統領広島へ 安倍首相の受け止めは？
ANN	2016/5/12 13:40	「招致はクリーンな形」官房長官"買収疑惑"否定
	2016/5/12 8:01	小泉進次郎氏らが厚労省分割案「役所 1 つでは困難」
	2016/5/11 23:30	【報ステ】ALS 患者が参考人出席できず
JNN	2016/5/12 6:35	政府 破壊装置命令を解除、PAC3 など撤収
	2016/5/12 6:35	ALS 患者の質疑拒否、自民・民進 主張に食い違い
	2016/5/11 19:22	政府後押しも・・・USJ 沖縄進出を見送り
FNN	2016/5/12 15:50	舛添知事「政治資金」問題 菅長官「知事自身が説明されると」
	2016/5/12 13:59	北方領土交渉 安倍首相、新たなアプローチは「脱・歴史」か
	2016/5/12 9:41	民進・山尾政調会長、自身の選挙区内の有権者に花代や香典支出

こうした選択基準により、2015 年 12 月 15 日から 2016 年 3 月 31 日までの期間において、271 の政治関連ニュース動画を収集した<sup>11</sup>。国会審議映像と同様に、音声データを抽出し、音声認識にかける一方、音声データの反訳からニュース原稿を作成した。音声認識による文字情報とニュース原稿を比較し、任意の文字列が 2 文字だけ続いた文字列 (bigram) の一致度を「正文率」とすると、全体としての平均は 77.53 % となった (表 1)。また、図 8 は正文率の分布を図示している。部分的な分析であるが、国会審議映像の音声認識による正文率が約 85% であることを考慮すると<sup>12</sup>、政治関連のニュース動画の音声認識はそれを下回ることになる。

ニュース動画に関する音声認識の相違を体系的に解明するため、ここでは正文率を従属変数とする多変量解析を試みた。分析対象である動画には政治関連のニュースが広範に含まれることを考慮すると、国会の会議録に基づく音声認識では、失言、失態、不祥事に関するニュースは適切に把握できないのかも知れない。一方、政治的な問題に焦点があるニュースの場合には正文率は

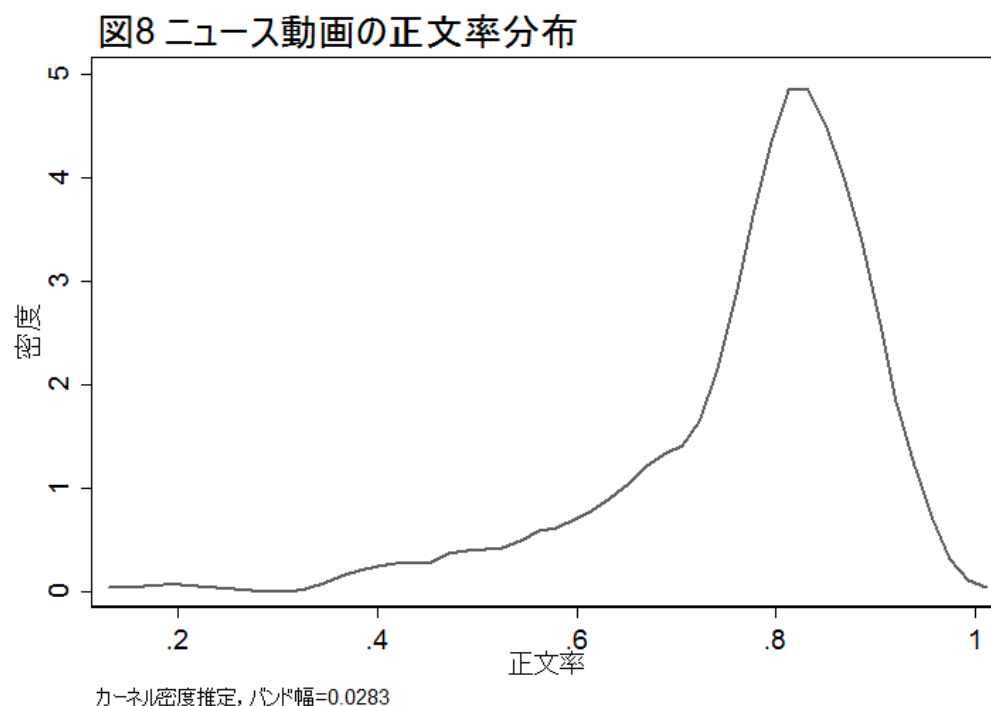
<sup>11</sup> 参考資料として、テレビ系列毎に、配信日、タイトル、再生時間、正文率を整理している。再生時間は平均 84 秒 (最小 21, 最大 378, 標準偏差 57) である。

<sup>12</sup> 第 190 回通常国会 (2016 年 1 月 4 日～6 月 1 日) における衆参両院の本会議、予算委員会、法務委員会、総務委員会、財政金融委員会、農林水産委員会を対象とした冒頭 10 分までの分析に拠れば (会議数 231), 正文率は平均 84.38, 最小 64.6, 最大 95.8, 標準偏差 5.64 である (Masuyama 2016)。

比較的に高く、また多岐にわたる国内の問題よりも国際的な外交・安全保障問題は内容的に均質であり、正文率も高いことが予測される。さらに、地理的、地域的に限定されるニュースのほう  
がそうでない一般的なニュースよりも内容的な均質性は高いであろう。報道時間の長いニュース  
ほど内容は多岐にわたり、均質性が低まる一方、報道で同じ言葉が繰り返されるほど内容の均質  
性は高いことが期待される。

表 1：政治関連ニュース動画の音声認識正文率

	平均	最小	最大	標準偏差	個数
全体	77.53	15.90	98.20	12.70	271
NNN	76.82	15.90	98.20	14.13	88
ANN	76.68	40.00	93.30	12.37	61
JNN	77.61	43.90	92.90	12.59	49
FNN	79.04	37.50	94.90	11.27	73



こうした正文率と諸要因の関連性を検証するにあたって、0～1 の範囲に制約された比率を従属変数とすることから、一般化線形モデル（二項分布・対数オッズ変換）の最尤法推定に拠ることとした（表 2）。ニュースのタイトル分類から、「不祥事」は政治家の失言や失態に関する報道とみなされるものを示すダミー変数とし、「外交安保」は日本の対外関係や安全保障に関する報道とみなされるものを示すダミー変数としている。また、報道内容のテキスト分析から頻出単語を抽出し、各ニュース中の最頻出単語の分類から、「政治」は選挙や大臣といった政治に関する単語を示すダミー変数とし、「地域」はアメリカや沖縄といった国名や地名を示すダミー変数としている。それぞれの推定係数は不祥事-0.216, 外交安保 0.203, 政治 0.166, 地域 0.251 であり、



失言や失態に関する報道であると正文率は相対的に低いが、外交安保に関する報道であると正文率は比較的に高くなり、政治に関する単語や地域に関する単語の頻度が高いほど、正文率も高いという傾向が確認される。

表 2：音声認識正文率の一般化線形モデル推定

	係数	標準誤差*	有意性
不祥事	-0.216	0.104	0.037
外交安保	0.203	0.122	0.095
政治	0.166	0.081	0.041
地域	0.251	0.125	0.044
対数再生時間	-0.448	0.156	0.004
対数最頻出数	0.321	0.192	0.094
定数	2.586	0.493	0.000

N271. 対数疑似尤度 -100.974. \* Huber-White 推定.

残差逸脱度 20.070. Pearson 逸脱度 21.524.

さらに、「対数再生時間」は動画の再生時間を対数化した数値であり、その推定係数からは再生時間が長いほど、正文率は低くなる関係にあることがわかる（-0.448）。また、「対数最頻出数」は最頻出単語の出現回数を対数化した数値であり、その推定係数は特定の単語の出現回数が多いほど、正文率は比較的に高くなることを示している（0.321）。こうした推定結果に拠ると、ニュース動画の音声認識による正文率は報道内容の均質性に規定されていると言えるようである。

ただし、推定係数の効果は直観的にわかり易いものではなく、正文率の変化に換算し直して確認する必要がある。例えば、不祥事に関する報道であるかないかに関しては、再生時間を平均値の約 83.7 秒、最頻出数を平均値の約 5.6 回とすると、0.761 と 0.719 が予測値となり、約 4 % の引き下げ効果があると言える。また、再生時間については、最頻出数を平均値とした場合、60 秒、90 秒、120 秒のそれぞれで 0.787、0.755、0.730 が予測値となり、逡減的な引き下げ効果のあることがわかる<sup>13</sup>。

#### 4. おわりに

本稿では、文字情報に偏ってきた国会審議や政治報道に対する従来の分析から脱し、音声や画像、映像を活用する試みとして、国会審議について議員の発言内容に対応する審議映像を検索し、

<sup>13</sup> 独立変数間の分散拡大要因は最大 2.41 であり、平均 1.70 である。ニュース配信元をダミー変数として追加した推定も行ったが、いずれも統計的に有意な効果を及ぼすものではなく、配信元による相違は確認できなかった。正文率の実測値と予測値の相関は 0.361 である。各独立変数の基礎統計、推定モデルの逸脱度は参考資料に整理している。図 8 に示すように、正文率の分布はやや偏ったものであり、補対数対数変換による推定も行ったが、表 2 の推定結果と大差ないものであった。

該当する審議映像の部分的視聴を可能にする「国会審議映像検索システム」を概説し、音声認識によって映像情報と文字情報を同刻する技術をニュース動画に適用する可能性を検証してきた。

国会審議映像検索システムは、国会会議録の文字情報から、発言に対応する映像にピンポイントで到達することを目指すものであり、その瞬間の審議映像の部分再生から、会議録では分からない発言者の表情や臨場感、会議の流れなどを把握できるようにする。また、審議映像に字幕を付すことにより、聴覚障害者や視覚障害者が国会審議映像を利用することも可能になり、発言の瞬間を URL として表現とすることにより、SNS を通じて審議映像をインターネットで共有することも容易になる。さらに、国会審議には会議録に含まれない様々な情報があり、例えば、委員会審議で用いられる参考資料についても、国会審議映像検索システムを活用すれば、部分再生した画像から瞬時に確認することができる。

国会審議映像の検索を会議録の公開前に音声認識による文字情報から行う試みは、審議映像検索の速報的な利用を促進するだけでなく、会議録といった「正解」のない映像検索に道を開くものである。本稿では、YAHOO!JAPAN ニュースの「映像」から政治関連のニュース動画を収集し、それらの動画から抽出した音声データに基づいて、音声認識による文字情報の「正解率」を分析した。ニュース動画に関する音声認識の精度は平均で約 78% となり、それは国会審議映像の音声認識を下回るものの、報道内容が国会だけに限られず、多岐にわたることを反映しているようである。本稿の分析は試論的なものであるが、国会会議録では捨象される立法情報の多角的な時空間の解明に寄与する第一歩と言えよう。

## 参考文献

- 秋田佑哉・河原達也. 2013. 「音声認識を用いたオンライン自動字幕作成・編集システム」『日本音響学会秋季研究発表会講演論文集』2-8-4.
- 河原達也. 2012. 「議会の会議録作成のための音声認識一衆議院のシステムの概要」情報処理学会研究報告 SLP-93-5.
- Masuyama, Mikitaka. 2012. “Text-based Search on Diet Deliberation Video Clips” *Presented at the 2012 Annual Meeting of the Association for Asian Studies*, Toronto, Canada.
- Masuyama, Mikitaka. 2016. “Reducing the Space and Time between Citizens and Parliament: Video Retrieval System for Diet Deliberations” *Prepared for delivery at the 2016 General Conference of the European Consortium for Political Research*, Charles University, Prague, Czech Republic.
- Masuyama, Mikitaka and Kaori Takeda. 2014. “Instant Parliamentary Deliberations Are in Our Reach” *Presented at the 2014 Annual Meeting of the American Political Science Association*, Washington, D.C., USA.
- 増山幹高. 2016. 「国会審議映像検索システムとそのニュース動画への応用可能性」政策研究大学院大学.
- 増山幹高・竹田香織. 2012. 「いかに見たい国会審議映像に到達するか？」2012 年度日本選挙学会（筑波大学）.
- 鈴木泰山・内山雄司・青木保一・相良毅・秋田佑哉・河原達也・竹田香織・増山幹高. 2014. 「音



## 参考資料

### (1) 基礎統計

	平均	標準偏差	最小	最大
正文率	0.775	0.127	0.159	0.982
不祥事	0.269	0.444	0	1
外交安保	0.177	0.382	0	1
政治	0.421	0.495	0	1
地域	0.100	0.300	0	1
再生時間	83.679	56.596	21	378
最頻出数	5.598	2.983	1	22
対数再生時間	4.282	0.499	3.045	5.935
対数最頻出数	1.608	0.470	0	3.091

### (2) 逸脱度

$$\text{残差逸脱度 } \sum_j d_j^2 \quad d_j^2 = 2 \left( y_j \log \frac{y_j}{\mu_j} + (1 - y_j) \log \frac{1 - y_j}{1 - \mu_j} \right)$$

$$\text{Pearson 逸脱度 } \sum_j r_j^2 \quad r_j^2 = \frac{(y_j - \mu_j)^2}{\mu_j(1 - \mu_j)}$$

### (3) 補対数対数変換による推定

	係数	標準誤差*	有意性
不祥事	-0.114	0.056	0.043
外交安保	0.102	0.058	0.078
政治	0.087	0.042	0.040
地域	0.121	0.061	0.047
対数再生時間	-0.228	0.077	0.003
対数最頻出数	0.160	0.093	0.084
定数	1.087	0.244	0.000

N271. 対数疑似尤度 -100.999. \* Huber-White 推定.

残差逸脱度 20.119. Pearson 逸脱度 21.573.